MATHÉMATIQUES ET SCIENCES HUMAINES

BRUNO LECLERC

Sur le nombre d'éléments des niveaux des produits de chaînes et des treillis permutoèdres

Mathématiques et sciences humaines, tome 112 (1990), p. 37-48

<http://www.numdam.org/item?id=MSH_1990__112__37_0>

Math. Inf. Sci. hum., (28e année, n°112, *1990*, p.37-48)

SUR LE NOMBRE D'ÉLÉMENTS DES NIVEAUX DES PRODUITS DE CHAÎNES ET DES TREILLIS PERMUTOÈDRES

Bruno LECLERC[1]

RÉSUMÉ - *Les produits de chaînes comptent parmi les ensembles (partiellement) ordonnés les plus fréquemment rencontrés. On rappelle, avec des démonstrations en partie nouvelles, divers résultats exacts ou approchés sur les cardinaux de leurs niveaux et sur le nombre de ses niveaux de cardinal maximum. Un plongement avec de bonnes propriétés permet d'appliquer ces résultats aux niveaux du permutoèdre (ordre faible de Bruhat sur les permutations).*

ABSTRACT - On the cardinalities of the levels of chain products and permutohedron lattices.
Cartesian products of chains are among the most frequently considered (partially) ordered sets. Several exact or asymptotic results on the cardinalities of their levels, and on the number of their maximum cardinality levels, are recalled. Several new proofs are proposed. An embedding with good properties allows us to apply these results to the levels of the permutohedron (the so-called weak Bruhat order on permutations).

1. INTRODUCTION

Pour c entier quelconque non négatif, on note $\underline{c+1}$ la chaîne (ensemble totalement ordonné) $\{0 < 1 < 2 < \ldots < c\}$ à c+1 éléments, qui seront parfois assimilés à des nombres entiers. Le produit de m chaînes $P = (\underline{c_1+1}) \times (\underline{c_2+1}) \times \ldots \times (\underline{c_m+1})$ est le produit cartésien des chaînes $\underline{c_1+1}, \ldots, \underline{c_m+1}$. On suppose sans perte de généralité que chacune de ces chaînes a au moins deux éléments : $c_i \geq 1$ pour $i = 1, \ldots, m$. Soit $x = (x_1,x_2,\ldots,x_m)$ un élément de P, avec $x_i \in \underline{c_i+1}$; on pose $r(x) = \sum_i x_i$. P est partiellement ordonné selon l'ordre produit usuel : pour deux éléments $x = (x_1,x_2,\ldots,x_m)$, $y = (y_1,y_2,\ldots,y_m)$ de P, on a $x \leq y$ si et seulement si $x_i \leq y_i$, pour $i = 1, \ldots, m$.

Les produits de chaînes tiennent une place importante parmi les ensembles (partiellement) ordonnés finis couramment considérés ; donnons deux exemples simples, mais bien différents, de situations où ils apparaissent :
- l'évaluation de candidats ou d'alternatives selon plusieurs critères totalement ordonnés revient à situer ces candidats dans un tel produit ;
- pour un nombre entier $p = p_1^{c_1} p_2^{c_2} \ldots p_m^{c_m}$, où $p_1, \ldots, p_m$ sont les facteurs premiers de p ; le treillis des diviseurs de p, munis de l'ordre de divisibilité, est isomorphe au produit de m chaînes $P = (\underline{c_1+1}) \times (\underline{c_2+1}) \times \ldots \times (\underline{c_m+1})$.

[1] CAMS-EHESS, 54 bd Raspail 75270 PARIS CEDEX 06

Les résultats concernant ces ensembles ordonnés revêtent donc en général un grand intérêt. La simplicité de leur définition leur confère une structure régulière, beaucoup plus simple *a priori* que, par exemple, celle du treillis des partitions, ou celle du treillis des ordres totaux (treillis permutoèdre). Ils constituent en fait une des généralisations les plus immédiates des treillis booléens, dont ils conservent bon nombre de propriétés.

La figure 1 ci-dessous donne les diagrammes de Hasse des produits $\underline{3}\times\underline{4}$ (a), $\underline{2}\times\underline{3}\times\underline{4}$ (b) et $\underline{2}\times\underline{2}\times\underline{5}$ (c).

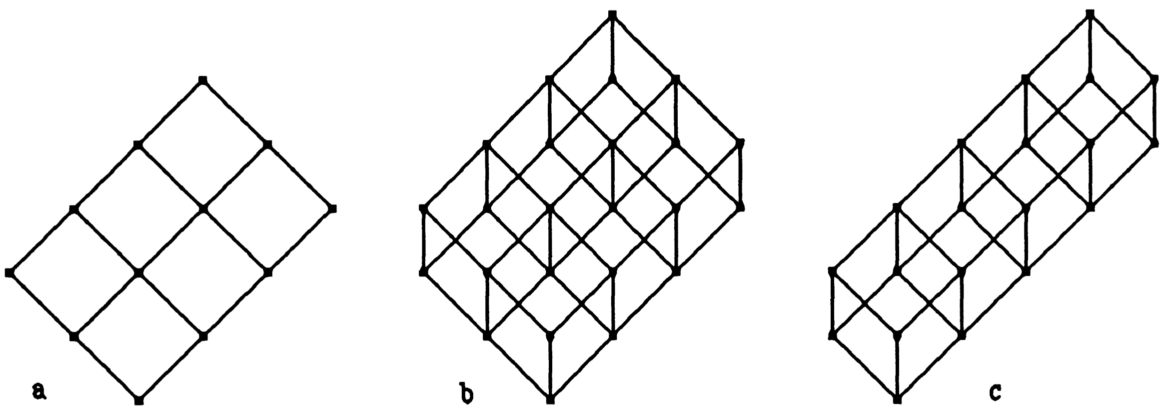

Figure 1

L'étude d'un produit de chaînes P, et notamment l'obtention de ses paramètres, n'est cependant pas toujours immédiate. Appelons *antichaîne* de P tout sous-ensemble A de P dont les éléments sont deux à deux incomparables, et soit $\alpha(P)$ la *largeur* de P, c'est-à-dire le cardinal maximum d'une antichaîne ; c'est aussi, d'après le théorème de Dilworth, le nombre minimum de chaînes requises pour partitionner P. Pour tout entier t, on définit de même une *t-antichaîne* comme un sous-ensemble A_t de P dont au plus t éléments sont deux à deux comparables et le *nombre de Dilworth* $\alpha_t(P)$ comme étant le cardinal maximum d'une t-antichaîne.

Rappelons qu'un ensemble ordonné P est dit *rangé* s'il existe une fonction de *rang* entière non négative r sur P telle que $r(x) = 0$ pour au moins un élément (minimal) de P, et que, si x couvre y dans P, alors $r(x) = r(y) + 1$ (on dit que x couvre y si $y \le z < x$ entraîne $y = z$). Il est immédiat que si P est un produit de chaînes $(\underline{c_1+1}) \times (\underline{c_2+1}) \times \ldots \times (\underline{c_m+1})$, la fonction r définie plus haut possède bien ces propriétés : x couvre y si toutes les coordonnées de x et de y sont égales, sauf une pour laquelle on a $x_i = y_i + 1$. Pour P rangé et $x \in P$, $r(x)$ est le *rang* de x ; la *hauteur* h, ou $h(P)$, de P est alors le rang maximum d'un élément de P ; par exemple, pour un produit de chaînes, on a $h = \sum_{1\le i\le m} c_i$. Pour tout entier $k \in [0,h]$, le *niveau* P_k de P est l'ensemble des éléments x de P tels que $r(x) = k$. On note n_k le nombre d'éléments de P_k, et on pose $n_k = 0$ pour tout entier $k \notin [0,h]$; on pose aussi $\nu = \max_k n_k$.

Une construction de De Bruijn et al. (1951) permet de voir qu'un produit de chaînes P est un ensemble ordonné "fortement de Sperner", c'est-à-dire que $\alpha_t(P)$ est égal au cardinal de la réunion des t plus grands niveaux de P ; en particulier, $\alpha = \nu$. La détermination des paramètres $\alpha_t(P)$ se ramène donc à celle des cardinaux des niveaux. Toutefois, le calcul de ces cardinaux, qui se pose dans bien des contextes (statistique, analyse hiérarchique...), n'est pas immédiat ; ainsi, Lerman donne dans son livre de 1981 une formule particulièrement compliquée et qui est, en fait, peu utilisable en pratique.

Un problème voisin est celui du dénombrement des niveaux de cardinal maximum (pour les exemples de la figure 1a, b, c, il y a respectivement 2, 1, et 3 tels niveaux) ; sa solution n'a été donnée qu'assez tardivement, et avec des démonstrations assez sophistiquées (Griggs 1984).

Le propos de cet article est de présenter ou rappeler quelques résultats, peut-être insuffisamment connus, sur la distribution, exacte ou approchée, des cardinaux des niveaux dans les produits de chaînes, en proposant parfois de nouvelles approches pour leur obtention.

Notre point de départ est une récurrence élémentaire, bien connue, permettant le calcul des cardinaux des niveaux d'un produit de chaînes. En reprenant cette récurrence dans un cadre un peu plus général, on réobtient d'abord, par des moyens tout à fait élémentaires, l'essentiel des résultats de Griggs mentionnés plus haut (théorème 5). On s'intéresse ensuite au calcul approché des cardinaux des niveaux lorsque le nombre de chaînes du produit augmente. On applique un théorème "limite-central" bien choisi pour donner une condition nécessaire et suffisante particulièrement simple pour que la distribution des cardinaux des niveaux tende vers la loi normale (théorème 6). On retrouve alors une formule asymptotique due à Anderson pour le calcul de α.

Un plongement canonique du treillis permutoèdre dans le produit de chaînes $\underline{2} \times \underline{3} \times \ldots \times \underline{m}$, établi par Le Conte de Poly-Barbut (1990 - ce numéro), permet d'appliquer également ces résultats à cet autre ensemble ordonné.

Le lecteur intéressé pourra trouver un grand nombre d'autres résultats sur les produits de chaînes dans les livres de Engel et Gronau (1985, chapitre 3) et de Anderson (1987, chapitre 4).

2. PRODUIT DIRECT D'UN ENSEMBLE ORDONNÉ ET D'UNE CHAÎNE.

Dans ce paragraphe, nous considérons la situation, plus générale que celle de l'introduction, où l'ensemble ordonné $P = P' \times (\underline{c+1})$ est le produit direct d'un ensemble ordonné P' et de la chaîne $(\underline{c+1})$. Nous allons établir que certaines propriétés de P' se transmettent alors à P. Un élément de P est noté $x = (x',j)$, avec $x' \in P'$ et $j \in \underline{c_m+1}$; on retrouve le cas précédent avec $P' = (\underline{c_1+1}) \times \ldots \times (\underline{c_{m-1}+1})$.

Tout d'abord, supposons que P' est rangé, de fonction de rang r' ; alors, P est aussi rangé, sa fonction de rang r étant donnée par $r(x) = r((x',j)) = r'(x') + j$ et son niveau P'_k étant l'ensemble $\{x' \in P' : r'(x') = k\}$; la hauteur h de P est donc $h' + c$, h' étant la hauteur de P'. Soit n'_k le cardinal de P'_k ; on pose $n'_k = 0$ pour tout entier $k \notin [0,h']$.

PROPOSITION 1 (Folklore). Les nombres n_k s'obtiennent en fonction des nombres n'_k par l'une ou l'autre des récurrences équivalentes :

$$n_k = \Sigma_{0 \le i \le c}\, n'_{k-i} \qquad (1)$$
$$n_{k+1} = n_k + n'_{k+1} - n'_{k-c}. \qquad (2)$$

Preuve. (1) est immédiat à partir de $r((x',j)) = r'(x) + j$; (2) en découle.□

Une interprétation de ce résultat est de dire que le nombre n_k est une "somme mobile" au sens de la moyenne mobile de la statistique : c'est la somme de $c + 1$ termes consécutifs (certains pouvant être nuls) parmi les n'_i.

On dira que P est *unimodal* (quant aux rangs) si la suite finie $n_0, n_1, n_2, \ldots, n_h$ est la concaténation d'une suite non décroissante $n_0, \ldots, n_p$ et d'une suite non croissante $n_{p+1}, \ldots, n_h$, et que P est *symétrique* (quant aux rangs) si on a $n_0 = n_h \leq n_1 = n_{h-1} \leq \ldots \leq n_k = n_{h-k} \leq \ldots$ (la symétrie ainsi définie implique l'unimodalité). Dès que la propriété d'unimodalité est vérifiée par P, les niveaux de P de cardinal v sont consécutifs ; on note q le nombre de ces niveaux ; plus généralement, pour tout ensemble ordonné rangé R, $q(R)$ sera le nombre de niveaux de R de cardinal maximum $v(R)$. On pose $v' = \max_k n'_k = v(P')$ et soit $q' = q(P')$ le nombre de niveaux de P' de cardinal v'.

PROPOSITION 2. Si P' est unimodal (resp. symétrique), alors P est unimodal (resp. symétrique).

Preuve. La suite des n_k commence par être croissante : $n_0 = n'_0$; $n_1 - n_0 = n'_1 > 0$. Supposons que P' est unimodal et soit i pour lequel elle décroit, c'est-à-dire que l'on a $n_{i+1} - n_i < 0$, donc, d'après (2), $n'_{i-c} > n'_{i+1}$. Dans ce cas n'_{i+1} ne peut être que dans la partie décroissante de la suite des n'_k, d'où $n'_{i+2} \leq n'_{i+1}$.

Si n'_{i-c} est encore dans la partie croissante, on a $n_{i+2} - n_{i+1} = n'_{i+2} - n'_{i-c+1} \leq n'_{i+1} - n'_{i-c} < 0$. Sinon, on a $n'_{i-c+1} \geq n'_{i+2}$ et donc $n_{i+2} - n_{i+1} \leq 0$; ces dernières inégalités restant vraies pour les différences de rangs suivantes, la suite des n_k ne peut se remettre à croître.

Si P' est symétrique, on a : $n_k = \Sigma_{0 \leq i \leq c}\, n'_{k-i} = \Sigma_{0 \leq i \leq c}\, n'_{h'-c+i} = \Sigma_{0 \leq i \leq c}\, n'_{h-c-k+i} = \Sigma_{0 \leq j \leq c}\, n'_{h-k-j} = n_{h-k}$, ce qui établit la symétrie de P.□

On s'intéresse maintenant au nombre q des niveaux de P de cardinal v. Les deux propositions suivantes permettent de déterminer ce nombre, connaissant q' et c. La partie (i) de la proposition 4 généralise aussi l'un des résultats donnés dans le livre de Anderson.

PROPOSITION 3. Si P' est unimodal et si on a $c < q'$, alors P a exactement $q' - c$ niveaux de cardinal v ; on a de plus dans ce cas $v = (c+1)v'$.

Preuve. Soit j le plus petit entier tel que $n'_j = v'$; avec $c < q'$ et k tel que $j + c \leq k \leq j + q' - 1$, on a $n_k = \Sigma_{0 \leq i \leq c}\, n'_{k-i} = (c+1)v'$. A partir de la récurrence (1), il est clair que cette valeur est la plus grande que puisse atteindre le cardinal d'un niveau de P.□

PROPOSITION 4. On suppose que P' est symétrique et que, de plus, $n'_{k+1} = n'_k$ implique soit $n'_k = 0$, soit $n'_k = v'$. Alors :

(i) de même $n_{k+1} = n_k$ implique soit $n_k = 0$, soit $n_k = v$.

(ii) Si $c \geq q'$, P a exactement un (pour h pair) ou deux (pour h impair) niveaux de cardinal maximum v.

Preuve. Montrons d'abord la partie (i). D'après la récurrence (2), l'égalité $n_{k+1} = n_k$ implique $n'_{k+1} = n'_{k-c}$, et dans tous les cas :

si $n'_{k+1} = 0$, on a aussi $n'_i = 0$ pour tout i inférieur à k, et donc $n_{k+1} = n_k = 0$;

si $n'_{k+1} = n'_{k-c} = v'$, on a $n'_i = v'$ pour tout i compris entre $k - c$ et $k+1$, et donc $n_{k+1} = n_k = (c + 1)v' = v$.

Supposons que $n'_{k+1} = n'_k$ implique soit $n'_k = 0$, soit $n'_k = v'$. Outre les cas précédents, on ne peut avoir, quand $n_{k+1} = n_k$, que $k - c = h' - k - 1$. Ce qui signifie que $h = h' + c$ est impair et que l'on a $k = {}^{(h-1)}/_2$ et $k+1 = {}^{(h+1)}/_2$, ce qui, avec la symétrie, entraîne bien $n_{k+1} = n_k = v$.

Pour (ii), remarquons que, d'après la symétrie, s'il y a plusieurs niveaux de cardinal v, ils sont "centrés" autour de la valeur ${}^h/_2 = {}^{(h'+c)}/_2$: il y a donc un entier j tel que l'on a $j < {}^{(h'+c)}/_2$ et $j+1 \geq {}^{(h'+c)}/_2$, et $n_j = v$. D'où $n_j = n_{j+1}$, tandis que, du fait que $q' \leq c$, on ne peut avoir $n'_j = n'_{j+1} = v$. On applique alors la partie (i) précédente.□

3. CAS D'UN PRODUIT DE CHAÎNES.

Les résultats précédents s'appliquent immédiatement à un produit de m chaînes $P = (\underline{c_1+1}) \times (\underline{c_2+1}) \times \ldots \times (\underline{c_m+1})$ (on pose dans ce paragraphe, et seulement dans ce paragraphe, $c_1 \geq c_2 \geq \ldots \geq c_m$). Le niveau $(\underline{c_1+1})_k$ est de cardinal 1 si $0 \leq k \leq c_1$, et de cardinal 0 sinon ; la récurrence de la proposition 1 permet de calculer les cardinaux des niveaux de $(\underline{c_1+1}) \times (\underline{c_2+1})$, puis de ceux de $(\underline{c_1+1}) \times (\underline{c_2+1}) \times (\underline{c_3+1})$, et ainsi de suite... tant que les nombres m et c_i, $i = 1,\ldots, m$ ne sont pas trop grands. De même, les propositions 2 à 4 ont pour conséquence le résultat suivant :

THEORÈME 5 (Griggs 1984, Engel et Gronau 1985)). Soit le produit de chaînes $P = (\underline{c_1+1}) \times (\underline{c_2+1}) \times \ldots \times (\underline{c_m+1})$. Alors :

(i) P est symétrique.

(ii) $n_{k+1} = n_k$ implique soit $n_k = 0$, soit $n_k = v$.

(iii) Si $c_1+1 > \sum_{2\leq i\leq m} c_i$, le produit de chaînes P a exactement $q = c_1+1 - \sum_{2\leq i\leq m} c_i$ niveaux de cardinal $v = \prod_{2\leq i\leq m} (c_i+1)$; sinon, P a deux niveaux de cardinal v si h est impair et un seul niveau de cardinal v si h est pair.

Preuve. Pour $m = 1$, la chaîne $(\underline{c_1+1})$ est symétrique et vérifie la condition (ii) ci-dessus ; à partir de la proposition 2 et de la partie (i) de la proposition 4, on en déduit par récurrence sur m, en prenant $P' = (\underline{c_1+1}) \times (\underline{c_2+1}) \times \ldots \times (\underline{c_{m-1}+1})$, que P possède ces mêmes propriétés. Le cardinal maximum d'un niveau $v = \alpha$ est donc égal à $n_{h/2}$ si h est pair, et à $n_{(h\pm 1)/2}$ si h est impair.

Le nombre de niveaux de cardinal maximum de la chaîne $\underline{c_1+1}$ est $q(\underline{c_1+1}) = c_1+1 > c_2$; d'après la proposition 3, on a donc $q((\underline{c_1+1}) \times (\underline{c_2+1})) = c_1 + 1 - c_2$ et $v((\underline{c_1+1}) \times (\underline{c_2+1})) = c_2 + 1$.

De même, si $c_1 + 1 - c_2 > c_3$, c'est-à-dire si $c_1 + 1 > c_2 + c_3$, on a :

$q((\underline{c_1+1}) \times (\underline{c_2+1}) \times (\underline{c_3+1})) = c_1+1 - c_2 - c_3$, et :

$v((\underline{c_1+1}) \times (\underline{c_2+1}) \times (\underline{c_3+1})) = (c_2 + 1)(c_3 + 1)$,

et ainsi de suite tant que $c_1 + 1$ reste strictement supérieur à $\sum_{2\leq i\leq m} c_i$.

Mais si l'on a, par exemple, $c_1 + 1 \leq c_2 + c_3$, alors, d'après la partie (ii) de la proposition 4, $q((\underline{c_1+1}) \times (\underline{c_2+1}) \times (\underline{c_3+1})) \leq 2$, et on voit facilement que ceci entraîne $q(P) \leq 2$.□

Exemples :

1) pour $m = 3$, on peut développer complètement le calcul de v. On obtient :

$v = (c_2+1)(c_3+1) - K/4$, avec :

$K = 0$ si $c_1+1 > c_2+c_3$; $K = (c_3+c_2-c_1)(c_3+c_2-c_1+2)$ si $c_1+1 \leq c_2+c_3$ et si $c_3+c_2-c_1$ est pair ; $K = (c_3+c_2-c_1+1)^2$ si $c_1+1 \leq c_2+c_3$ et si $c_3+c_2-c_1$ est impair.

Évidemment, pour $m = 4$, la situation devient plus complexe...

2) Nous dirons que P est un *(m,c)-hypercube* (on dit aussi, dans un certain contexte, une *algèbre de Post*) si $c_1 = c_2 = \ldots = c_m = c$; pour $c = 1$, la récurrence (1) devient $n_{k+1} = n'_k + n'_{k+1}$: c'est celle des nombres binomiaux. Pour c quelconque, les cardinaux des niveaux du (m,c)-hypercube constituent une généralisation de ces nombres. Par exemple, pour $c = 5$, le cardinal n_k du niveau P_k correspond au nombre de façons d'obtenir le score $k+m$, en jetant m dés à jouer du modèle ordinaire. La table suivante donne les nombres n_k correspondants pour $1 \leq m \leq 4$:

m \ k	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	1	1	1	1	1	1															
2	1	2	3	4	5	6	5	4	3	2	1										
3	1	3	6	10	15	21	25	27	27	25	21	15	10	6	3	1					
4	1	4	10	20	35	56	80	104	125	140	146	140	125	104	80	56	35	20	10	4	1

Le (m,c)-hypercube possède deux niveaux de cardinal v si le produit mc est impair, un seul sinon.

3) Le produit de chaînes $\underline{2} \times \underline{3} \times \ldots \times \underline{m}$ a $m!$ éléments. La table suivante donne les nombres n_k correspondants pour $2 \leq m \leq 5$:

m \ k	0	1	2	3	4	5	6	7	8	9	10
2	1	1									
3	1	2	2	1							
4	1	3	5	6	5	3	1				
5	1	4	9	15	20	22	20	15	9	4	1

Certains lecteurs ont sans doute reconnu ce tableau de nombres, sur lequel nous reviendrons au paragraphe 6 ci-dessous.

4. APPROXIMATION PAR LA LOI NORMALE.

Les résultats précédents ont permis d'établir quelques propriétés de la suite des $n_k(P)$ et de proposer un calcul de ces nombres par récurrence, au moins pour m pas trop grand. Le calcul des probabilités fournit des outils efficaces pour l'approximation du nombre d'éléments de P de rang compris entre deux valeurs données.

On s'occupe ici de dénombrements et la probabilité que nous introduisons consiste simplement à considérer que tous les éléments de P sont *a priori* équiprobables. Lorsque x parcourt l'ensemble des éléments de P, chaque coordonnée x_i est donc vue comme une variable qui prend les valeurs 0, 1, ..., c_i avec la même fréquence $1/(c_i+1)$, indépendamment des autres variables x_j, $j \neq i$. La variable $r(x)$ se trouve donc être une somme de m variables indépendantes.

La moyenne μ_i et la variance V_i de la variable x_i sont celles de la loi discrète uniforme :

$$\mu_i = \sum_{k=0}^{c_i} \frac{k}{c_i+1} = \frac{c_i}{2}, \text{ et } V_i = \left(\sum_{k=0}^{c_i} \frac{k^2}{c_i+1}\right) - \left(\frac{c_i}{2}\right)^2 = \frac{c_i(c_i+2)}{12}. \tag{3}$$

La moyenne μ et, du fait de l'indépendance, la variance V de $r(x)$ s'obtiennent par sommation sur i ; nous notons s l'écart-type (racine carrée de la variance) de $r(x)$:

$$\mu = \frac{1}{2}\sum_{i=1}^{m} c_i = \frac{h}{2},\ V = \frac{1}{12}\sum_{i=1}^{m} c_i(c_i+2), \text{ et } s = \sqrt{V}. \tag{4}$$

Quand m augmente indéfiniment, la variable $^{(r(x)-\mu)}/_{s}$ tend vers la loi normale (ou loi de Laplace-Gauss), de moyenne 0 et de variance 1, pourvu que la suite des variables $(x_i)_{i=1,2,...}$ satisfasse les conditions de l'un des théorèmes "limite-centraux". En fait, nous pouvons donner une condition particulièrement simple, nécessaire et suffisante pour avoir cette convergence.

THEORÈME 6. Une condition nécessaire et suffisante pour que la distribution de la variable $r(x)$ tende vers une loi normale de moyenne μ et de variance V est :

$$\mathrm{Lim}_{m\to\infty} \frac{c_m}{s} = 0 \tag{5}$$

Preuve. Nous montrons essentiellement que (5) équivaut à la condition de Lindeberg (1922 ; cf. par exemple Feller 1950, Hennequin et Tortrat 1965 ou, pour l'énoncé seul, Rouanet et Leclerc 1970). Dans le cas discret qui nous intéresse, cette condition prend la forme suivante, dans laquelle la seconde sommation porte sur les valeurs de k :

pour tout $\varepsilon > 0$,
$$\mathrm{Lim}_{m\to\infty} \frac{1}{V}\sum_{i=1}^{m}\left(\sum_{\frac{c_i}{2}\geq|k-\mu_i|>\varepsilon s} \frac{(k-\mu_i)^2}{c_i+1}\right) = 0. \tag{6}$$

Dans le cas général, (6) équivaut à la conjonction de trois propriétés :

(i) la convergence vers la loi normale évoquée ci-dessus ;
(ii) $V \to \infty$;
(iii) $^{V_m}/_{V} \to 0$.

La propriété (ii) est vérifiée dans le cas qui nous intéresse puisque, du fait que $c_i \geq 1$ pour tout i, V croît indéfiniment avec m. La propriété (iii) est immédiatement équivalente à (5), qui est donc une conséquence de (6).

Réciproquement, si (5) est vérifiée, nous pouvons prendre, pour ε quelconque, m_0 suffisamment grand pour avoir $^{c_i}/_{s} \leq 2\varepsilon$, donc $\varepsilon s \geq {}^{c_i}/_{2}$, pour tout $i \geq m_0$. Alors, la condition (6) est vérifiée, car elle porte sur une expression où la sommation sur i ne prend en compte, pour $m \geq m_0$, qu'un nombre fini et décroissant avec m de termes non nuls.

Enfin, il est connu en théorie de l'addition des variables aléatoires qu'en l'absence de la propriété (iii) ci-dessus, on ne peut avoir de convergence vers la loi normale de la variable $r(x)$ que si la variable x_m tend elle-même vers la loi normale, ce qui n'est, par hypothèse, pas vrai ici.□

Donnons une interprétation géométrique simple à la condition (5) : on associe au produit de chaînes P un parallèlotope rectangle à m dimensions dont les longueurs des côtés sont les nombres c_i ; alors (5) équivaut à ce que le rapport de la longueur de tout côté à celle, égale à $\sqrt{\Sigma_i \; c_i^2}$, de la diagonale du parallélotope, tende vers zéro.
Le cas le plus simple où cette condition est vérifiée est celui où où les c_i sont bornés (pour tout i, $c_i < c$ fixé). Une autre situation sera considérée au paragraphe 6.

5. CALCULS APPROCHÉS.

Rappelons que la loi normale est définie par la densité $f(\xi) = \frac{1}{\sqrt{2\pi}} e^{-\frac{\xi^2}{2}}$. Cette fonction atteint son maximum, égal à $\frac{1}{\sqrt{2\pi}}$, pour $\xi = 0$. En tenant compte du paramètre d'échelle s, et du cardinal n de P, on retrouve une formule asymptotique pour v due à Anderson (1967) :

$$v \approx \frac{n}{s\sqrt{2\pi}} = \sqrt{\frac{6}{\pi}} \frac{\Pi_i \; (c_i + 1)}{\sqrt{\Sigma_i \; c_i(c_i + 2)}} \tag{7}$$

Dans le cas du (m,c)-hypercube, on obtient la forme particulière (8) ci-dessous :

$$v \approx \frac{\sqrt{6}(c+1)^m}{\sqrt{\pi m c(c+2)}} \tag{8}$$

Soit $\Phi(u) = \int_{-\infty}^{u} f(\xi) d\xi$ la fonction de répartition de la loi normale, et soient j et j' deux entiers tels que $0 \le j \le j' \le h$. On approche le nombre d'éléments de P contenus dans les niveaux j à j', inclus, par la quantité :

$$|\cup_{j \le i \le j'} P_k| \approx n[\Phi(\frac{j' + 0,5 - \mu}{s}) - \Phi(\frac{j - 0,5 - \mu}{s})]. \tag{9}$$

En particulier, on obtient ainsi une approximation du nombre $\alpha_t(P)$. Par exemple, pour h pair et t impair, on prend $j = {}^{(2h-t-1)}/_2$ et $j' = {}^{(2h+t+1)}/_2$ (le lecteur trouvera aisément les valeurs de j et de j' correspondant aux autres cas).

Pour $t = 1$, on obtient ainsi une approximation, par intervalles cette fois, de $v = \alpha$; elle est en principe plus précise que celle donnée par (7). Si P est tel qu'aucun des c_i n'est supérieur ou égal à la somme de tous les autres, on obtient, si h est pair, c'est-à-dire si seul le niveau $h/2$ est de cardinal maximum v :

$$v \approx n[\Phi(\frac{1}{2s}) - \Phi(- \frac{1}{2s})], \tag{10}$$

et si h est impair, c'est-à-dire si les niveaux ${}^{(h+1)}/_2$ et ${}^{(h-1)}/_2$ sont ceux de cardinal maximum v :

$$v \approx (n/2)[\Phi(\frac{1}{s}) - \Phi(-\frac{1}{s})]. \qquad (10')$$

Quand P a plus de deux niveaux de cardinal v, l'usage de l'approximation normale n'est pas raisonnable, puisque cet usage est lié à la propriété (5) ; d'après le théorème 5, si $q > 2$, le maximum des c_i ne peut être considéré comme petit devant la somme des autres.

Exemples. Illustrons l'emploi de l'approximation normale en donnant, dans deux cas particuliers, les valeurs exactes de v obtenues à partir de la récurrence (2) et les valeurs approchées obtenues à partir des formules (7),ou (8), et (10) ; on a utilisé la table de $\Phi(u)$ du recueil de Abramowitz et Stegun (1965).

Exemple 1)

P est le (4,5)-hypercube : $n = 1296$, $h = 20$; on a $v = 146$.

$v \approx 151$ par (8) [l'"erreur" est d'à peu près 3,4%] et $v \approx 150$ par (10) [2,7%].

Exemple 2)

$m = 5$; $c_1 = c_2 = 3$, $c_3 = c_4 = 5$, $c_5 = 6$, d'où $n = 4032$ et $h = 22$; on a $v = 440$.

$v \approx 458$ par (7) [4,1%] et $v \approx 455$ par (10) [3,4%].

Comme on pouvait s'y attendre, l'approximation normale est meilleure dans le premier cas, où les nombres c_i sont égaux ; dans les deux cas, le paramètre m est encore trop petit pour qu'elle puisse être très bonne.

6. APPLICATION AU PERMUTOÈDRE.

Le treillis permutoèdre S_m est un sous-ordre couvrant, et conservant les mêmes niveaux, du produit de chaînes $\underline{2} \times \underline{3} \times \ldots \times \underline{m}$ (Le Conte de Poly-Barbut 1990 - ce numéro). La figure 2 ci-dessous illustre cette propriété : le tracé du diagramme de Hasse de S_4 peut y être comparé à celui du produit de chaînes $\underline{2} \times \underline{3} \times \underline{4}$ donné par la figure 1(b).

Les cardinaux des niveaux du permutoèdre S_m sont donc les mêmes que ceux du produit de chaînes correspondant. Ainsi, une formule de récurrence donnée par Kendall (1962 ; cf. aussi Guilbaud et Rosenstiehl 1963) pour le calcul de ces cardinaux se retrouve en particularisant la proposition 1 au cas où $P = P' \times \underline{m}$ avec $P' = \underline{2} \times \underline{3} \times \ldots \times (\underline{m-1})$.

On sait que S_2 et S_3 ont deux niveaux de cardinal maximum (et S_4 un seul). Alors, pour $m \geq 3$, les inégalités $m - 1 \geq 2 \geq q(S_{m-1})$ permettent de montrer par récurrence, en utilisant la proposition 4, que S_m n'a pas plus de deux niveaux de cardinal maximum.

Les résultats du paragraphe 4 permettent de retrouver simplement des propriétés connues du coefficient τ de corrélation des rangs de Kendall. Du fait des fortes symétries du permutoèdre, la distribution de $\tau(O_1,O_2)$, où O_1 et O_2 sont deux ordres totaux pris indépendamment au hasard équiprobable, se ramène à celle de $\tau(O,O_1)$, où O est l'élément minimum du treillis permutoèdre (l'ordre total de référence) et O_1 un ordre total pris au hasard équiprobable, et l'on a : $\tau(O,O_1) = 1 - \dfrac{4r(O_1)}{m(m-1)}$.

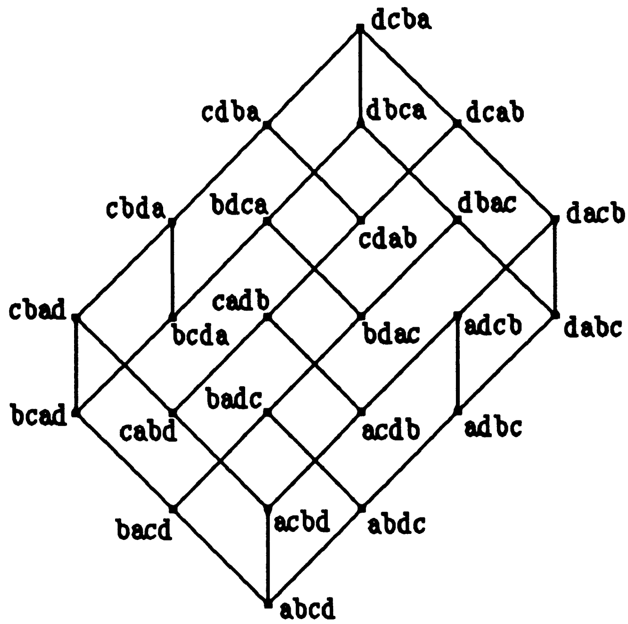

Figure 2

La moyenne μ et la variance V de la variable $r(O_1)$, rang de O_1 dans le treillis permutoèdre S_m, s'obtiennent alors immédiatement par particularisation des formules (4) ci-dessus ; on retrouve des expressions connues pour μ et pour V (dans l'article de Guilbaud et Rosenstiehl, une coquille s'est glissée dans l'expression de la variance) :

$$\mu = \frac{1}{2}\sum_{i=1}^{m} i = \frac{m(m+1)}{4}, \text{ et } V = \frac{1}{12}\sum_{i=1}^{m} i(i+2) = \frac{m(m+1)(2m+7)}{72}. \tag{11}$$

L'expression (11) de la variance a deux conséquences. D'abord, la forme asymptotique (7) pour le cardinal maximum d'un niveau prend l'expression particulière (12) suivante :

$$\nu \approx \frac{(m+1)!}{\sqrt{\pi m(m+1)(2m+7)}} \tag{12}$$

Ensuite, il est intéressant de noter qu'elle entraîne que la condition du théorème 6 est satisfaite par le produit de chaînes $\underline{2} \times \underline{3} \times \ldots \times \underline{m+1}$, et ce bien que la suite des c_i, qui est exactement celle des m premiers entiers 1, 2, ..., m, n'y soit pas bornée. Nous venons donc d'obtenir une démonstration tout à fait différente de celle de Kendall (op.cit.) de la convergence vers la loi normale de la distribution du coefficient τ de corrélation des rangs (sous hypothèse "nulle", c'est-à-dire d'équiprobabilité), un fait qui est à la base d'un test non paramétrique parmi les plus connus.

Terminons avec un exemple analogue à ceux de la fin du paragraphe 5 :
$m = 6$; $c_i = i$, pour $i = 1, 2, \ldots, 6$ (niveaux du permutoèdre S_7), d'où $n = 7! = 5040$, $h = 21$ et $\nu = 573$; on obtient :
$\nu \approx 630$ par l'approximation (8) [9,9% d'erreur] et $\nu \approx 619$ par l'approximation plus fine (9') [8,0%].

Ce cas où $c_i = i$ est assez défavorable, car les nombres c_i ne peuvent guère être considérés comme proches entre eux. On trouve dans les manuels de statistique non paramétrique des tables pour le calcul du coefficient τ jusqu'à $m = 40$, tables équivalentes, on l'a vu plus haut, à celles de la distribution exacte des cardinaux des niveaux du permutoèdre S_m: au delà de cette valeur, l'approximation normale est considérée en statistique comme largement suffisante en pratique.

7. CONCLUSION

Comme il a été dit plus haut, si les résultats donnés ci-dessus ne sont pour la plupart pas nouveaux, nous avons pu proposer pour plusieurs d'entre eux des démonstrations qui nous paraissent originales. Ainsi, la démonstration de Griggs du théorème 5 ci-dessus repose sur le fait que les produits de chaînes ont une propriété, classique en théorie des ensembles ordonnés mais assez sophistiquée, dite propriété "LYM" (pour laquelle on pourra à nouveau se reporter aux livres de Anderson et de Engel et Gronau, ou à celui de Barthélemy *et al.* 1991). La preuve de Engel et Gronau repose, elle, sur une récurrence sur m, mais nous semble néanmoins moins élémentaire que celle donnée ici. Notons que Griggs montre aussi que, pour $q \leq 2$, P est strictement de Sperner, c'est-à-dire que les seules antichaînes de cardinal maximum sont des niveaux, un résultat que Engel et Gronau attribuent de leur côté à Clements (1968).

Dans son livre, Anderson précise que la formule asymptotique (8) s'obtient "by using complex variable methods". C'est bien cette démarche que l'on a reprise en utilisant le théorème de Lindeberg, qui se montre à partir des fonctions caractéristiques (transformées de Fourier). Ce dernier n'est sans doute pas le plus connu des théorèmes limite-centraux et il nous paraît intéressant d'en signaler une application directe, à la rencontre du calcul des probabilités et de la théorie des ensembles ordonnés.

Par ailleurs, le plongement du permutoèdre dans un produit de chaînes paraît avoir un intérêt certain. Il rend par exemple le calcul de la variance du coefficient τ tout à fait immédiat. Retrouver la convergence vers la loi normale de la distribution des niveaux est moins élémentaire, puisque nous devons recourir au théorème de Lindeberg. Mais on peut noter que pour montrer cette convergence, Kendall utilise de son côté un autre résultat fort, celui selon lequel la convergence des moments entraîne celle en loi.

Notons que nombre de problèmes résolus dans les produits de chaînes ont pour analogues dans le permutoèdre des questions ouvertes. Le problème le plus connu est de savoir si S_m est ou non un ordre de Sperner, c'est-à-dire si l'égalité $\alpha(S_m) = \nu(S_m)$ est vraie pour m quelconque (cf. Griggs 1988). On trouve en fait peu de renseignements sur l'état de cette question dans la littérature ; nous avons pu vérifier à la main que S_5 est partitionnable en chaînes symétriques, donc fortement de Sperner (la même constatation est immédiate pour S_3 et facile pour S_4). On peut de la même façon poser d'autres questions telle que la suivante : le cardinal maximum d'une antichaîne A de S_m "sans compléments", c'est-à-dire ne pouvant contenir simultanément un ordre O_1 et son inverse est-il toujours égal à $n_{(h/2)\pm 1}$ si h est pair et à $n_{(h\pm 1)/2} = \nu$ si h est pair ?

BIBLIOGRAPHIE

ABRAMOWITZ M., STEGUN I. A. (1965), *Handbook of Mathematical Functions*, Dover, New York.
ANDERSON I. (1967), On primitive sequences, *J. London Math. Soc.* 42, 137-148.
ANDERSON I. (1987), *Combinatorics of Finite sets*, Clarendon Press, Oxford.
BARTHÉLEMY J.P., LECLERC B., MONJARDET B. (1991), *Les ensembles ordonnés*, en préparation.
CLEMENTS G.F. (1968), A generalization of Sperner's theorem on subsets of a finite set, non publié ; cf. *Notices of the Amer. Math. Soc.* 16 (1969), 700.
DE BRUIJN N.G., TENGBERGEN C., KRUYSWIJK D. (1951), On the set of divisors of a number, *Nieuw Arch. Wiskd.* 23, 191-3.
ENGEL K., GRONAU H.D. (1985), *Sperner Theory in Partially Ordered Sets*, Teubner, Leipzig.
FELLER W. (1950), *An Introduction to Probability Theory and Its Applications*, vol. 1, 2nd edition, Wiley, New York.
GRIGGS J.R. (1984), Maximum antichains in the product of chains, *Order* 1, 21-28.
GRIGGS J.R. (1988) Problems on chain partitions, *Discrete Math.* 72, 157-162.
GUILBAUD G. Th., ROSENSTIEHL P. (1963), Analyse algébrique d'un scrutin, *Math. Sci. hum.* 4, 9-33.
HENNEQUIN P.L., TORTRAT A. (1965), *Théorie des probabilités et quelques applications*, Masson, Paris.
KENDALL M. G. (1962), *Rank Correlation Methods*, 3rd ed., Hafner, New York.
LE CONTE de POLY-BARBUT C. (1990), Le diagramme du permutoèdre est intersection des diagrammes de deux produits d'ordres totaux, *Math. Inf. Sci. hum.* 112, 1990, 49-53.
LERMAN I. C. (1981), *Classification et analyse ordinale des données*, Dunod, Paris.
LINDEBERG J. W. (1922), Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung, *Mathematische Zeitschrift* 15, 211-225.
ROUANET H., LECLERC B. (1970), Le rôle de la distribution normale en statistique, *Math. Sci. hum.* 32, 57-74.